

XII. TECHNICAL NOTES

Disclosure-avoidance procedures

Like other statistical agencies, the Bureau of Economic Analysis (BEA) is legally required to safeguard the confidentiality of the information that it receives. In addition, like other agencies, it must balance its responsibility to avoid disclosing confidential information with its responsibility to release as much information as possible. It balances these responsibilities by presenting the estimates for regions, states, and local areas only at the North American Industry Classification System (NAICS) subsector level or Standard Industrial Classification (SIC) two-digit level, even though it receives source data at the NAICS four- and five-digit industry levels or SIC three- and four-digit levels.

Most of the data series that BEA receives from other agencies are not confidential. The agencies summarize their data by program, county, or state, so that each record, or data cell, contains data for enough individuals or establishments to preclude the identification of data for a specific individual or establishment and, therefore, to preclude disclosure of confidential information.¹

However, the Quarterly Census of Employment and Wages (QCEW, formerly known as ES-202 data) tabulations that BEA receives from the Bureau of Labor Statistics (BLS) include records that would disclose confidential information. The confidential information on wages and salaries for some business firms is identifiable from the state and county estimates of wages and salaries that are derived from the QCEW data.²

To prevent either the direct or the indirect disclosure of the confidential information, BEA uses the BLS state and county nondisclosure file. BEA uses as many BLS nondisclosure cells as possible, but cannot use some of them for various reasons. The most important reasons are that the industry structure published by BEA does not exactly match the NAICS subsector or SIC two-digit detail provided by BLS and that BEA does not use QCEW data for the farm sector. When BEA drops BLS nondisclosure cells, other cells must be selected to prevent the disclosure of confidential information. In order to determine which estimates should be suppressed, the total wages and salaries file and the wages-and-salaries-nondisclosure file are used to prepare a multidimensional matrix. This matrix is tested, and the estimates that should be suppressed are selected.³

Dual allocation

Dual allocation is a statistical procedure that forces the elements of a matrix to sum to column and row control totals. It is used to adjust, for instance, a preliminary estimate of income by state and industry so that sum of income in

¹ For a list of some of the agencies that provide data to BEA, see “Sources of the data” in the Chapter II Sources and Methods.

² For specific information, see Chapter III Wage and Salary Disbursements.

³ In this test, computer programs impose a set of rules and priorities on this matrix so that the estimates that should be suppressed are selected until indirect disclosure is impossible.

an industry across all states equals a national control total for that industry and simultaneously the sum of income in a state across all industries equals a control total for that state. It is also used to adjust a preliminary estimate of quarterly state personal income so that it is consistent with both national control totals by quarter and annual state control totals.

Specifically, dual allocation subtracts the sum of the algebraic values in a row from the row control total. It divides this difference by the sum of the absolute values in the row and then multiplies the resulting ratio by the absolute value of each element in the row and adds the result to the algebraic value of that element. This procedure is repeated for each row and then a parallel procedure is repeated for each column. The whole process is repeated five times.

After the fifth repetition, any differences between the row and column control totals and the output matrix row and column sums are eliminated by a process called feathering. This is accomplished by selecting the first column with a non-zero difference and the first non-zero row difference with the same sign. The smaller of the two differences is subtracted from the element in that row and column and from the final row and column sums. This procedure forces the difference between either the final row sum and its corresponding control total or the final column sum and its corresponding control total to zero.

Before performing any subtraction, the element in the row and column selected is checked for a zero value and to see if the subtraction would cause a change in the element's sign. If either of these tests is true, the next non-zero row difference with like sign is selected.

The entire feathering process is repeated until all differences between final column sums and column control totals have been forced to zero. At this point the row sums and row control totals will also be equal.

Employment

BEA gives equal weight to full-time and part time jobs in its estimates of employment. Wage and salary jobs and proprietors' jobs are counted, but unpaid family workers and volunteers are not. Proprietors' employment consists of the number of sole proprietorships and the number of partners in partnerships. Wage and salary employment is on a place of work basis. Proprietors' employment, however, is more nearly by place of residence because, for nonfarm sole proprietorships, the estimates are based on IRS tax data that reflect the address from which the proprietor's individual tax return is filed, which is usually the proprietor's residence. Nonfarm partnership employment reflects the tax-filing address of the partnership, which may be either the residence of one of the partners or the business address of the partnership.

The employment estimates are designed to be consistent with the estimates of wage and salary disbursements, proprietors' income, and earnings. The employment estimates are based on the same sets of source data as the corresponding earnings estimates and are prepared with parallel methodologies. Two components of proprietors' income—the income of limited partnerships and

the income of tax-exempt cooperatives—have no corresponding employment estimates.

Imputation

One of the principles of the national income and product accounts (NIPAs) is that they reflect market transactions. In a few instances, a comprehensive account of total income and production requires BEA to impute a value or a transaction. This keeps the NIPA invariant to how certain activities are carried out. For instance, some transactions, such as the provision of food, lodging, and clothing to employees have an element of barter—food is bartered for labor (at least in part). In this case, imputation involves placing a market value on the food employees received so that the estimate of their total compensation is comprehensive and invariant to changes in the proportions received in cash and in kind. In other transactions, such as the rental of housing to an owner-occupant, no transaction appears in the records of the economy. In this case, imputation involves constructing a transaction between a producer and a consumer (who happen to be the same person) and placing a market value on the housing services exchanged. If the imputation were not made, then housing output and consumption would fall if a household purchased the house it had been renting.

The imputations described here are those that affect state personal income.⁴ They are: pay-in-kind, employer-paid health and life insurance premiums, the value of food and fuel produced and consumed on farms, the net rental value of owner-occupied housing, the net margins on owner-built housing, the value of depositor services furnished without payment by financial intermediaries except life insurance carriers, premium supplements for property and casualty insurance, and the interest received from life insurance carriers. These imputations accounted for about 8 percent of personal income at the national level in 2003.

Imputed pay-in-kind is added to the estimates of wage and salary disbursements so that all the earnings of employees who receive part of their wages in pay-in-kind will be included in personal income. This imputation is an estimate of the value of the food, lodging, clothing, and other goods and services that are received by employees from their employers as partial or full payment for their services.

The imputation for employer-paid health and life insurance premiums is included in employer contributions for employee pension and insurance funds, a component of supplements to wages and salaries.

The imputed value of food and fuel produced and consumed on farms is included in farm proprietors' income so that that measure reflects the income from all of the production of noncorporate farms.

The imputed net rental value of owner-occupied housing is included in the rental income of persons. The imputation assumes that the owner-occupants are

⁴ See table 7.12, "Imputations in the National Income and Product Accounts," *Survey of Current Business* 84 (February 2004): 24-5.

in the rental business and that they are renting the houses in which they live to themselves: As tenants, they pay rent to the landlords (that is, to themselves); as landlords, they collect rent from their tenants (that is, from themselves), they incur expenses, and they may have a profit or a loss from the rental business.

The imputed net margin on owner-built housing is included in proprietors' income, classified in the construction industry. It represents the net income of individuals from the management of the construction or renovation of their own dwellings and is included in the measure of the output of structures.

The imputed value of depositor services furnished without payment by financial intermediaries except life insurance carriers is included in personal interest income. The value of depositor services is received by persons from depository institutions, that is, from commercial banks, mutual savings banks, savings and loan associations, credit unions, and regulated investment companies. It is an estimate of the value of the services (such as checking and record keeping) that these institutions provide to persons without an explicit charge.⁵

The premium supplement for property and casualty insurance is the imputed value of financial protection and intermediation services insurance companies provide to policyholders.⁶ It is included in personal interest income.

Also included in personal interest income is the imputed interest received from life insurance carriers. It consists of the property income earned on life insurance and pension reserves. This income is attributed to policyholders in order to include it in personal saving, rather than in business saving, and when the income is earned, rather than when it is distributed.

Industrial classification

For the private sector the North American Industry Classification System (NAICS) is used for the industrial classification of wage and salary disbursements, employer contributions for employee pension and insurance funds, and proprietors' income. NAICS is used for 2001 to the present at the state level and from 2000 to the present at the county level.⁷

For earlier years the Standard Industrial Classification (SIC) was used. The *Standard Industrial Classification Manual, 1967* was used for the years 1969-74, the 1972 *Manual* was used for the years 1975-87, and the 1987 *Manual* was used for 1988-2001 for states and 1988-2000 for counties.⁸

For the public sector, the estimates of wages and salaries and employer contributions for employee pension and insurance funds are classified by level of

⁵ See "Measuring the Services of Commercial Banks in the NIPAs," *Survey 83* (September 2003):33-44.

⁶ See "Measuring the Services of Property-Casualty Insurance in the NIPAs," *Survey 83* (October 2003): 10-26.

⁷ Office of Management and Budget: *North American Industry Classification System, United States, 2002* (Lanham, MD: Bernan Press, 2002)

⁸ Executive Office of the President, Office of Management and Budget, Statistical Policy Division, *Standard Industrial Classification Manual, 1967* (Washington, DC: U.S. Government Printing Office (GPO), 1967); *Manual, 1972* (GPO, 1972); *Manual, 1987* (GPO, 1987).

government—Federal, state, and local. The estimates for the Federal government are subclassified into civilian and military.

Interpolation and extrapolation

Interpolation and extrapolation are used to prepare the first approximations of some components of personal income for the years in which direct source data are unavailable. Both procedures use the data for these components for benchmark years—the years for which the best data are available—and both frequently use other data that are related to the benchmark-year data for the components.

Interpolation is used to derive the first approximation of estimates for years that are between benchmark years. For example, if data for wages and salaries for an industry were available only from the decennial censuses of population but employment data were available annually from another source, the first approximations of wages and salaries for 1981-89 could be interpolated from the wages and salaries data for 1980 and 1990, the two census benchmark years, and from the employment data for 1980-90.

Extrapolation is used to derive first approximations for years that are beyond the most recent benchmark year. For example, the first approximations of wages for 1991-99 might be extrapolated from the census benchmark data for 1990 and from the employment data for 1990-99. The estimates based on extrapolation are usually superseded by revised estimates when benchmark data become available for a more current year. For the preceding example, the estimates for 1991-99 would be superseded by estimates based on interpolation when census benchmark data became available for 2000.

Both interpolation and extrapolation are illustrated in the following examples. In the first two examples, interpolation is used to derive the first approximations of wages and salaries for an industry in areas A, B, and C for the years 2 and 3 that are between the benchmark years 1 and 4. In the third example, extrapolation is used to derive the approximations for year 5.

In the first example, “straight-line interpolation” is used to derive the first approximations for years 2 and 3 from the data for benchmark years 1 and 4.⁹ The first approximations for year 2 equals the amount for year 1 plus one-third of the increase from year 1 to year 4; the preliminary estimate for year 3 equals the amount for year 1 plus two-thirds of the increase.

Wages and salaries in thousands of dollars

| | Year 1 (benchmark) | Year 2 (interpolation) | Year 3 (interpolation) | Year 4 (benchmark) |
|--------------|-----------------------|---------------------------|---------------------------|-----------------------|
| Area A | 28 | 34 | 40 | 46 |

⁹ Straight-line interpolation assumes that the magnitude of the annual change is the same in each year in the interpolated time series, subject to modification by the adjustment to the national control totals. Straight-line interpolation is used as the default option, when no annual source data related to the income series are available.

| | | | | |
|--------------|----|----|----|----|
| Area B | 34 | 43 | 53 | 62 |
| Area C | 74 | 81 | 87 | 94 |

In the second example, interpolation with a related series of data, the indicator series, is used to derive the first approximations for years 2 and 3 from the benchmark data for years 1 and 4 and from the indicator series for all four years. The data for wages and salaries are the benchmark data, the employment data are the indicator series, and the average wages (computed as wages and salaries divided by employment) are the interpolation ratios.¹⁰ This method of interpolation is illustrated in three steps.

First, average wages for years 1 and 4 are calculated from the wage and employment data for those years. Wages for each year are divided by the number of employees for the year to yield the average wages of the employees.

Employment and average wages

| | Year 1 | | Year 4 | |
|--------------|------------|-----------------------------|------------|-----------------------------|
| | Employment | Average Wages in dollars | Employment | Average wages in dollars |
| Area A | 4 | 7,000 | 4 | 11,500 |
| Area B | 6 | 5,667 | 10 | 6,200 |
| Area C | 11 | 6,727 | 10 | 9,400 |

Second, straight-line interpolation is used to derive average wages for years 2 and 3 from average wages for years 1 and 4.

Average wages in dollars

| | Year 1 (benchmark) | Year 2 (interpolation) | Year 3 (interpolation) | Year 4 (benchmark) |
|--------------|-----------------------|---------------------------|---------------------------|-----------------------|
| Area A | 7,000 | 8,500 | 10,000 | 11,500 |
| Area B | 5,667 | 5,845 | 6,022 | 6,200 |
| Area C | 6,727 | 7,618 | 8,509 | 9,400 |

Third, the interpolated average wages for each year are multiplied by the employment data for each year to yield the first approximations.

¹⁰ Using an indicator series for interpolation between two benchmark years assumes that any change in the relationship between the data for the income component for the benchmark years and the data from the indicator series for the benchmark years occurs uniformly over time. This relationship is embodied in the interpolation ratios, which in this example are the average wages. For this procedure, straight-line interpolation of the benchmark-year interpolation ratios is used to calculate the ratios for the intervening years. A benchmark-year interpolation ratio is the ratio of the datum for an income component for the benchmark year to the datum for the same year from the annual indicator series. The interpolation ratios for the intervening years are multiplied by the data for those years from the indicator series to yield the interpolated series for those years.

Employment and wage approximations

| | Year 2 | | Year 3 | |
|--------------|-----------------------|------------|-----------------------|------------|
| | Wages in thousands | | Wages in thousands | |
| | Employment | of dollars | Employment | of dollars |
| Area A | 5 | 43 | 4 | 40 |
| Area B | 7 | 41 | 9 | 54 |
| Area C..... | 10 | 76 | 9 | 77 |

In the third example, extrapolation with an indicator series is used to derive the first approximations of wages for year 5 from average wages for year 4—used here as the extrapolation ratios—and employment data for year 5.¹¹ The average wages are multiplied by employment to yield the first approximations of wages for year 5.

First approximations of wages for year 5

| | Year 4 | Year 5 | |
|-------------|--------------------------------|-----------------------|------------|
| | Average Wages in dollars | Wages in thousands | |
| | | Employment | of dollars |
| Area A..... | 11,500 | 5 | 58 |
| Area B..... | 6,200 | 12 | 74 |
| Area C..... | 9,400 | 9 | 85 |

After interpolation or extrapolation is used to calculate the first approximations of a component of personal income, the approximations are adjusted proportionately to sum to the component's control total.

Per capita personal income

Per capita personal income is calculated as the personal income of the residents of a given area divided by the resident population of that area. In computing per capita personal income for states and counties, BEA uses the Census Bureau's annual midyear population estimates. Except for college student and other seasonal populations, which are measured on April 1, the population for all years is estimated on July 1.

Local area per capita personal income estimates should be used with caution for several reasons. In some instances, an unusually high or low per capita personal income is the temporary result of unusual conditions, such as a bumper crop or hurricane. In other instances, the income levels of certain groups atypical of the resident population may cause a longer term high or low per capita

¹¹ Using an indicator series for extrapolation assumes that the relationship between the data for the income component for the latest benchmark year and the data from the indicator series for that year remains unchanged in the subsequent years.

personal income that is not indicative of the economic well-being of the area. For instance, a major construction project—such as a defense facility, power plant, or dam—may substantially raise the per capita personal income of an area for several years because it attracts highly paid workers whose income is measured at the construction site. This high per capita income is not indicative of the economic well-being of the permanent residents of the area (or, in many cases, of the resident construction workers themselves, because they frequently send a substantial portion of their wages to their dependents living in other areas).

Conversely, the presence of a large institutional population—such as that of a college or a prison—will tend to keep the per capita personal income of an area at a lower level because the residents of these institutions have little income attributable to them at these institutions. This lower per capita personal income is not indicative of the economic well-being of the other residents of the area (or, in some cases, of the institutional populations, because some of these populations, such as college students, typically receive support from their families living in other areas).

The per capita personal income estimates can also be misleading in areas where population changes rapidly. Population is measured at midyear, whereas income is measured as a flow over the year; therefore, a significant change in the population of an area during the year, particularly if it occurs around midyear, can cause a distortion in the per capita personal income estimates.

In counties where farm income predominates, additional considerations should be taken into account. Farm proprietors' income as measured for personal income reflects returns from current production; it does not measure current cash flows. Sales out of inventories are included in current gross farm income, but they are excluded from net farm income because they represent income from a previous year's production. Additions to inventories are included in net farm income at current market prices; therefore, farmers' attempts to regulate their cash flows by adjusting inventories are not reflected in BEA's farm proprietors' income estimates. However, this regulation of cash flows by farmers extends their earnings cycles, so it helps them to survive losses or lowered income for 2 or 3 years. In addition, the per capita personal income of sparsely populated counties that are dependent on farming will react more sharply to vagarious weather, world market demand, and changing government policies related to agriculture than that of counties where the sources of income are more diversified.

Personal income, adjusted gross income, and money income

The measure of personal income that is prepared by BEA differs substantially from adjusted gross income (AGI), which is the principal measure of the income of individuals that is tabulated by the Internal Revenue Service. Personal income also differs from money income, which is prepared by the Census Bureau.

As compared with AGI, personal income consists of the income of nonprofit institutions serving individuals, private noninsured welfare funds, and

private trust funds as well as of individuals, whereas AGI consists only of the income of individuals who file individual income tax returns. Personal income also includes employer contributions to private health and pension funds and to government employee retirement plans, several types of imputed incomes, transfer payments, and all of the interest received by individuals, whereas AGI excludes all employer contributions, imputed incomes, most transfer payments, and the nontaxable interest received by individuals. Personal income, unlike AGI, excludes personal contributions for social insurance, realized capital gains and losses, and pensions and annuities from private and government employee retirement plans.¹²

Personal income differs from money income mainly because money income consists only of the income that is received by individuals in cash and its equivalents. Personal income, unlike money income, includes imputed income, lump-sum payments not received as part of earnings, certain in-kind personal current transfer receipts—such as Medicaid, Medicare, and food stamps—and employer contributions to private health and pension funds and to government employee retirement plans. Personal income, unlike money income, excludes personal contributions for social insurance, pensions and annuities from private and government employee retirement plans, and income from interpersonal transfers, such as child support.

Personal income for a given area and year includes the income received by individuals living in that area during that year. In contrast, money income for a given area and year consists of the income received during the year by individuals living in the area on April 1 of the following year, regardless of where they were living when they received the income. The income received by individuals who died or moved abroad before April 1 of the following year is not included in the money income of any area.

Personal income is prepared quarterly for states and annually for counties, whereas money income for states, counties, and cities is prepared decennially on the basis of data from the “long-form” sample of the census of population.¹³

¹² For more information, see Mark A. Ledbetter, “Comparison of BEA Estimates of Personal Income and IRS Estimates of Adjusted Gross Income,” *Survey* 84 (April 2004): 8-22.

¹³ The Small Area Income and Poverty Estimates program of the Census Bureau has prepared post-censal estimates of median household income for counties. In addition, the Census Bureau prepares estimates of median household income for states using data from the annual Current Population Survey.